⬤ *MOTOROLA*

FAX TRANSMITTAL SHEET

Motorola, Inc.
Intellectual Property Section
Law Department
1303 E. Algonquin Rd.
Schaumburg, IL 60196

Telephone: (847) 576-3635
Facsimile: (847) 576-3750

|48| **Number of Pages (including this page)**

Date: October 10, 2005

To: Examiner Shingles, K. – Group 2141

Location: United States Patent and Trademark Office

Fax No.:   571-273-8300

From: Steven A. May (Registration No. 44,912)

Subject: Serial No. 09/447,400 –Chen et al.

MESSAGE:

Enclosed herewith, please find an APPEAL BRIEF for filing in the below-identified application.

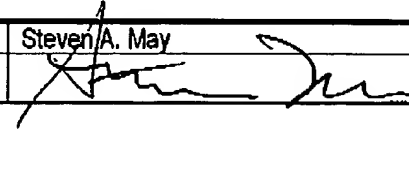## PLEASE GIVE THESE PAPERS TO:

EXAMINER: **Shingles, K.**
GROUP ART UNIT: 2141
SERIAL NO.:   09/447,400
FILED: 11/22/1999
INVENTOR: Chen et al.
ATTORNEY DOCKET NO.:   CE08051R

PTO/SB/21 (08-00)

TRANSMITTAL FORM (to be used for all correspondence after initial filing)	Application Number	09/447,400
	Filing Date	November 22, 1999
	First Named Inventor	Chen et al.
	Group Art Unit	2141
	Examiner Name	Shingles, K.
Total Number of Pages in this Submission	Attorney Docket Number	CE08051R

ENCLOSURES (check all that apply)

☐ Fee Transmittal Form	☐ Assignment Papers (for an Application)	☐ After Allowance Communication to Group
☐ Fee Attached	☐ Drawing(s)	☐ Appeal Communication to Board of Appeals and Interferences
[X] Appeal Brief	☐ Licensing-Related papers	☐ Appeal Communication to Group (Appeal Notice, Brief, Reply Brief)
☐ After Final	☐ Petition	☐ Proprietary Information
☐ Affidavits/Declaration(s)	☐ Petition to Convert to a Provisional Application	☐ Status Letter with appropriate copies
☐ Extension of time Request	☐ Power of Attorney, Revocation, Change of Correspondence Address	☐ Other Enclosure(s) (please identify below) ☐ Response to Restriction Requirement ☐ Associate Power of Attorney ☐ RCE
☐ Express Abandonment Request		
☐ Information Disclosure Statement	☐ Terminal Disclaimer	☐ Copy of Notice to File Missing Parts ☐ Transmittal of Formal Drawings
☐ Certified Copy of Priority Documents	☐ Request for Refund	☐ Response to Notice of Non-Recordation of Document
☐ Response to Missing Parts/ Incomplete Application	☐ CD, Number of CDs	
☐ Response to Missing Parts Under 37 CFR 1.52 or 1.53	Remarks	

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm or Individual	Steven A. May	Registration No.	44,912
Signature			
Date	October 10, 2005		

CERTIFICATE OF TRANSMISSION

I hereby certify that this correspondence is being facsimile transmitted to the USPTO to facsimile number 571-273-8300 or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450 on the date listed below:

| Typed or printed name | Nanette Orr | | |
| Signature | | Date | October 10, 2005 |

FEE TRANSMITTAL

Patent fees are subject to annual revision

[ ] Applicant claims small entity status. See 37 CFR 1.27

	Complete if Known
Application Number	09/447,400
Filing Date	November 22, 1999
First Named Inventor	Chen et al.
Examiner Name	Shingles, K.
Group Art Unit	2141
Attorney Docket No.	CE08051R

RECEIVED CENTRAL FAX CENTER
OCT 1 0 2005

TOTAL AMOUNT OF PAYMENT	($)

METHOD OF PAYMENT (check all that apply)

[ ] Check [ ] Credit card [ ] Money Order [ ] Other [ ] None

[X] Deposit Account:

Deposit Account Number: **502117**

Deposit Account Name: **Motorola, Inc.**

The Director is authorized to: (check all that apply)

[X] Charge fee(s) indicated below [X] Credit any overpayments

[ ] Charge any additional fee(s) during the pendency of this application

[ ] Charge fees(s) indicated below, except for the filing fee to the above-identified deposit account.

FEE CALCULATION

1. BASIC FILING FEE

Large Fee Code	Entity Fee ($)	Small Fee Code	Entity Fee ($)		Fee Paid
1001	770	2001	385	Utility filing fee	
1002	340	2002	170	Design filing fee	
1003	530	2003	265	Plant filing fee	
1004	780	2004	385	Reissue filing fee	
1005	160	2005	80	Provisional filing fee	

SUBTOTAL (1) ($)

2. EXTRA CLAIM FEES

	Previously Paid**		Extra Claims		Fee from below		Fee Paid
Total Claims	20	-		=		X 50 =	
Independent Claims	10	-		=		X 200 =	
Multiple Dependent					280 =		

Large Fee Code	Entity Fee ($)	Small Fee Code	Entity Fee ($)	Fee Description
1202	18	2202	9	Claims in excess of 20
1201	84	2201	42	Independent claims in excess of 3
1203	280	2203	140	Multiple dependent claim, if not paid
1204	84	2204	42	* Reissue independent claims over original patent
1205	18	2205	9	*Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) ($)

**or number previously paid, if greater. For Reissues, see above.

FEE CALCULATION (continued)

3. ADDITIONAL FEES

Large Entity Fee Code	Fee ($)	Small Entity Fee Code	Fee ($)	Fee Description	
1051	130	2051	65	Surcharge – late filing fee or oath	
1052	50	2052	25	Surcharge – late Provisional filing	
1053	130	1053	130	Non-English specification	
1812	2520	1812	2520	For filing a request for ex parte Reexamination	
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	
1805	1840*	1805	1840*	Requesting publication of SIR after Examiner action	
1251	110	2251	55	Extension for reply within first month	
1252	420	2252	210	Extension for reply within second month	
1253	950	2253	475	Extension for reply within third month	
1254	1480	2254	740	Extension for reply within fourth month	
1255	2010	2255	1005	Extension for reply within fifth month	
1401	500	2401	165	Notice of Appeal	
1402	500	2402	165	Filing a brief in support of an appeal	500
1403	290	2403	145	Request for oral hearing	
1451	1510	1451	1510	Petition to institute a public use proceeding	
1452	110	2452	55	Petition to revive – unavoidable	
1453	1330	2453	665	Petition to revive – unintentional	
1501	1330	2501	665	Utility issue fee (or reissue)	
1502	480	2502	240	Design issue fee	
1503	640	2503	320	Plant issue fee	
1460	130	1460	130	Petitions to the Commissioner	
1807	60	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of IDS	
8021	40	8021	40	Recording each patent assignment per property (times number of properties)	
1809	770	2809	385	Filing a submission after final rejection (37 CFR § 1.129(a))	
1810	770	2810	385	For each additional invention to be examined (37 CFR § 1.129(b))	
1801	770	2801	385	Request for Continued Examination (RCE)	
1802	900	1802	900	Request for expedited examination of a design application	

Other fee (specify)

SUBTOTAL (3) ($)

SUBMITTED BY

Name (Print/Type)	Steven A. May		Complete (if applicable)		
		Registration No.	44,912	Telephone	847-576-3635
Signature		Date	October 10, 2005		

RECEIVED
CENTRAL FAX CENTER

- PATENT -

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

OCT 10 2005

APPLICANT:	Chen et al.	EXAMINER:	Shingles, K.
SERIAL NO.:	09/447,400	ART UNIT:	2141
FILED:	11/22/99	CASE NO.:	CE08051R
ENTITLED:	LOAD BALANCING METHOD IN A COMMUNICATION NETWORK		

Motorola, Inc.
Corporate Offices
1303 E. Algonquin Road
Schaumburg, IL 60196
October 10, 2005

APPEAL BRIEF UNDER 37 CFR 1.192

Certificate of Transmission under 37 CFR 1.8

I hereby certify that this correspondence is being
facsimile transmitted to the United States
Patent and Trademark Office.

on _____

Motorola, Inc.   10/10/2005

Name of applicant, assignee, or   Date
Registered Representative

_____
Signature

Mail Stop Appeal Brief - Patents
Commissioner of Patents
P.O. Box 1450
Alexandria, Va. 22313-1450

Commissioner:

     The appellants hereby respectfully submit the following Appeal Brief in response to a final Office Action dated May 10, 2005, and a Notice of Appeal filed August 8, 2005.

1. REAL PARTY IN INTEREST

The real party in interest in this appeal is Motorola, Inc.

2. RELATED APPEALS AND INTERFERENCES

There are no other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal.

3. STATUS OF CLAIMS

This is an appeal from a final Office Action, dated May 10, 2005. Claims 1-14 are appealed. In a first Office Action dated August 12, 2002, the Examiner rejected claims 1-4, 7, and 8 under 35 U.S.C. §103(a) as being unpatentable over Dias et al. (U.S. patent no. 6,119,143) in view of Smith (U.S. patent no. 5,835,724). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Adelman et al. (U.S. patent no. 6,006,259). The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Attanasio et al. (U.S. patent no. 5,918,017) and Fine (U.S. patent no. 4,894,846). The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of the applicants' admitted prior art. The Examiner rejected claims 10-14 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Yu (U.S. patent no. 6,078,943). In a response dated November 12, 2002, the appellants replied to the first Office Action without amending the claims.

In a second Office Action dated January 29, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu (U.S. patent no. 6,078,943) in view of Smith. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Smith and further in view of

2

the applicants' admitted prior art. In a response dated April 28, 2003, the appellants replied to the second Office Action without amending the claims.

In a third Office Action dated July 14, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt et al. (U.S. patent no. 6,424,992) and further in view of Devarakonda (U.S. patent no. 6,195,680). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated October 14, 2003, the appellants replied to the third Office Action without amending the claims.

In a fourth Office Action dated December 30, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and further in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated June 22, 2004, the appellants replied to the fourth Office Action without amending the claims.

In a fifth Office Action dated October 5, 2004, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda

3

and further in view of the applicants' admitted prior art.   In an Amendment dated February 4, 2005, the appellants amended each of claims 1-4.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T.

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2.

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method

4

including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers.

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period.

In a final Office Action dated May 10, 2005, the Examiner rejected claims 8 and 9 under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claims the subject matter which the applicant regards as the

5

invention, in particular as lacking an antecedent basis for the phrase "said combination." The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal et al. (U.S. patent no. 6,327,622, hereinafter referred to as "Jindal"). The Examiner rejected claims 11-14 under 35 U.S.C. §103(a) as being unpatentable over Jindal in view of Couland et al. (U.S. patent no. 6,253,230). No claims were allowed. The pending claims 1-14 are reproduced below in the attached Appendix.

4. STATUS OF AMENDMENTS

An Amendment and Response to the Final Office Action was filed on August 8, 2005, and is currently pending. In the Response to the Final Office Action, the appellants amended claims 8 and 9 and responded to the Examiner's rejection of claims 1-14. The appellants received an Advisory Action, dated September 1, 2005. The Advisory Action reiterated the rejections of claims 1-14 under the cited prior art.

5. SUMMARY OF INVENTION

The appellant's invention provides a method and apparatus for load balancing by assigning load in a communication system network having multiple servers, each server having a load level based on serving a number of clients.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. (FIG. 5; pages 19-20; pages 25-26)

6

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2. (FIG. 5; pages 20-21; pages 25-26)

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers. (FIG. 5; page 21; pages 25-26)

7

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period. (FIG. 5; page 22; pages 25-26)

6.   ISSUES

Whether claims 1-4 are unpatentable under 35 U.S.C. §35 U.S.C. §102(e) as being anticipated by Jindal.

7.   GROUPING OF CLAIMS

Appellants designate the following group of claims:
Group I: claims 1-14.

8.   ARGUMENT

(i) Rejection under 35 U.S.C. §112, first paragraph:

None

(ii) Rejection under 35 U.S.C. §112, second paragraph:

8

None

(iii) Rejection under 35 U.S.C. §102:

The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal. With respect to claim 1, the Examiner stated that Jindal teaches grouping multiple servers into first and second server groups, wherein said first server group has a load level less than a load level of said second server group (col. 9, line 48 to col. 10, line 40) calculating a time period T (col. 2 lines 40-67; col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27), assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; and col. 9, line 14-27), determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group (col. 8, line 55 to col. 10, line 40), and assigning load to a server selected from said another group of servers after expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27).

The appellants respectfully disagree with the Examiner's interpretation of Jindal. Jindal teaches a selection and subsequent re-selection of a server from a same group of servers. That is, Jindal teaches a central server that periodically determines a server from a same group of servers to be the "preferred" server to which client requests are routed. In order to select the preferred server, the central server periodically collects one or more items of information concerning each server in the group, such as a number of clients being serviced by the server at that moment, a number of client requests handled by the server in a predetermined period of time, and a distance of the server from the central server. Based on the collected information, the central server then selects a server from the group of servers to be the preferred server. The group of servers may comprise multiple servers directly connected to the central server (servers 110, 112, and 114 in FIG. 2), or the group of servers may comprise multiple servers that are indirectly connected to the central server (servers 302, 304, 312, and 314, which are connected to central server 100 via intermediate servers 306 and 316 in FIG. 3, that is, the 'server farms' referenced by the Examiner in the Advisory Action). However, in both instances,

9

at each server selection point the preferred server is selected by the central server from the same group of servers. Thus load balancing is performed not by varying the groups of servers analyzed but instead by periodically analyzing each server in a same group of servers.

In the Advisory Action, the Examiner referenced column 10, lines 32-40, of Jindal as teaching the segregation of servers into multiple groups. The appellants respectfully disagree. In this section of Jindal, Jindal merely teaches that the server group may comprise individual, participating servers ('segregated' servers) distributed among separate 'server farms,' rather than all the servers of each of one or more 'farms.' These individual servers may be considered something of a virtual group. Jindal then teaches that each farm, upon receiving a load allocation for the participating servers of that farm and which load allocation is based on a 'global' load distribution policy, may then distribute load to the participating servers of that farm based on a different load distribution policy than the 'global' load distribution policy. However, again, at each selection point, the load is being distributed among the same participating servers, that is, among the members of the virtual group.

By contrast, claim 1 teaches load balancing by varying the groups of servers considered. That is, claim 1 teaches grouping multiple servers into a first and second server groups, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of a time period T, determining another group of servers comprising the group of servers that includes the first server group and further comprising the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. Nowhere does Jindal teach such load balancing. Accordingly, the appellants respectfully submit that claim 1 is not unpatentable over the prior art of record.

Each of claims 2-4 includes limitations of assigning load to a server selected from a group of servers comprising a first server group from an initial time until expiration of said time period T, determining multiple other groups of servers comprising the group of servers that includes the first server group and further comprising another server group of multiple other servers groups, and assigning load to a server selected from a group of

10

servers of the multiple other groups of servers after expiration of said time period T. As noted above, none of these limitations are taught by Jindal. Accordingly, the appellants respectfully submit that claims 2-4 are not unpatentable over the prior art of record.

Regarding dependent claims 5-14, because claims 5-14 depend directly or indirectly from independent claim 4, the appellants respectfully submit that claims 5-14 are not unpatentable over the prior art of record.

(iv) Rejection under 35 U.S.C. §103:

None.

(v) Other rejections

None.

8.   CONCLUSION

For the above reasons, the appellants respectfully submit that the rejection of claims 1-14 under 35 U.S.C. §102(e) as being upatentable over Jindal is in error and should be reversed and the claims allowed.

Respectfully submitted,

David Chen, et al.

By: _____

Steven A. May
Attorney for Appellants
Registration No. 44,912
Tel. No.: 847/576-3635
Fax No.: 847/576-3750

11

APPENDIX

1. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a first and second server groups, wherein said first server group has a load level less than a load level of said second server group;

calculating a time period T;

assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T;

determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group; and

assigning load to a server selected from said another group of servers after expiration of said time period T.

2. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating time periods $T1$ and $T2$, wherein said time period $T2$ is longer than said time period $T1$;

assigning load to a server selected from a group of servers comprising said server group G0 from an initial time until expiration of said time period $T1$;

determining another group of servers comprising said group of servers that includes said server group G0 and further comprising said server group G1;

assigning load to a server selected from said another group of servers after expiration of said time period $T1$;

determining yet another group of servers comprising said group of servers that includes said another group of servers and further comprising said server group G2; and

12

assigning load to a server selected from said yet another group of servers after expiration of said time period T2.

3. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers.

4. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk corresponding to said server groups G1 through Gk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

13

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers, wherein each group of servers of said other groups of servers corresponds to an expiring time period.

5. The method as recited in claim 4 wherein said plurality of time periods T1 through Tk each is based on a difference between load levels of at least two server groups in said plurality of server groups G0 through Gk.

6. The method as recited in claim 4 further comprising the step of:

receiving an update of load level of at least one of said plurality of servers in said plurality of server groups G0 through Gk;

repeating said grouping to produce a new plurality of server groups G0 through Gk based on said update of load level;

repeating said calculating said plurality of time periods to produce a new plurality of time periods T1 through Tk corresponding to said new plurality of server groups G0 through Gk;

resetting said initial time to a reset initial time, and assigning load to a server selected from servers in said new server group G0 from said reset initial time until expiration of said new time period T1;

assigning load, after expiration of each of said new time periods T1 through Tk measured from said reset initial time, to a server selected from servers in a combination of servers including said new server group G0 and at least one other server group, in said new server groups G1 through Gk, corresponding to an expiring time period.

14

7. The method as recited in claim 4 wherein said grouping of said plurality of server groups G0 through Gk is based on similarity of load levels among said plurality of servers.

8. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a round robin selection method.

9. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a random selection method.

10. The method as recited in claim 4 wherein each of said plurality of time periods T1 through Tk is based on load levels of at least two server groups selected from said plurality of server groups G0 through Gk, a request arrival rate and a server service rate.

11. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of said plurality of servers.

12. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of a combination of servers of said plurality of servers.

13. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of said plurality of servers.

14. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of a combination of servers of said plurality of servers.

- PATENT -   **OCT 1 0 20**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT: Chen et al.

EXAMINER: Shingles, K.

SERIAL NO.: 09/447,400

ART UNIT: 2141

FILED: 11/22/99

CASE NO.: CE08051R

ENTITLED: LOAD BALANCING METHOD IN A COMMUNICATION
NETWORK

Motorola, Inc.
Corporate Offices
1303 E. Algonquin Road
Schaumburg, IL 60196
October 10, 2005

APPEAL BRIEF UNDER 37 CFR 1.192

Mail Stop Appeal Brief - Patents
Commissioner of Patents
P.O. Box 1450
Alexandria, Va. 22313-1450

Commissioner:

The appellants hereby respectfully submit the following Appeal Brief in response
to a final Office Action dated May 10, 2005, and a Notice of Appeal filed August 8, 2005.

1. REAL PARTY IN INTEREST

The real party in interest in this appeal is Motorola, Inc.

2. RELATED APPEALS AND INTERFERENCES

There are no other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal.

3. STATUS OF CLAIMS

This is an appeal from a final Office Action, dated May 10, 2005. Claims 1-14 are appealed. In a first Office Action dated August 12, 2002, the Examiner rejected claims 1-4, 7, and 8 under 35 U.S.C. §103(a) as being unpatentable over Dias et al. (U.S. patent no. 6,119,143) in view of Smith (U.S. patent no. 5,835,724). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Adelman et al. (U.S. patent no. 6,006,259). The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Attanasio et al. (U.S. patent no. 5,918,017) and Fine (U.S. patent no. 4,894,846). The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of the applicants' admitted prior art. The Examiner rejected claims 10-14 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Yu (U.S. patent no. 6,078,943). In a response dated November 12, 2002, the appellants replied to the first Office Action without amending the claims.

In a second Office Action dated January 29, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu (U.S. patent no. 6,078,943) in view of Smith. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Smith and further in view of

2

the applicants' admitted prior art. In a response dated April 28, 2003, the appellants replied to the second Office Action without amending the claims.

In a third Office Action dated July 14, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt et al. (U.S. patent no. 6,424,992) and further in view of Devarakonda (U.S. patent no. 6,195,680). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated October 14, 2003, the appellants replied to the third Office Action without amending the claims.

In a fourth Office Action dated December 30, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and further in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Attanasio and Fine The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated June 22, 2004, the appellants replied to the fourth Office Action without amending the claims.

In a fifth Office Action dated October 5, 2004, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda

3

and further in view of the applicants' admitted prior art. In an Amendment dated February 4, 2005, the appellants amended each of claims 1-4.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T.

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2.

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method

4

including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers.

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period.

In a final Office Action dated May 10, 2005, the Examiner rejected claims 8 and 9 under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claims the subject matter which the applicant regards as the

5

invention, in particular as lacking an antecedent basis for the phrase "said combination." The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal et al. (U.S. patent no. 6,327,622, hereinafter referred to as "Jindal"). The Examiner rejected claims 11-14 under 35 U.S.C. §103(a) as being unpatentable over Jindal in view of Couland et al. (U.S. patent no. 6,253,230). No claims were allowed. The pending claims 1-14 are reproduced below in the attached Appendix.

4. STATUS OF AMENDMENTS

An Amendment and Response to the Final Office Action was filed on August 8, 2005, and is currently pending. In the Response to the Final Office Action, the appellants amended claims 8 and 9 and responded to the Examiner's rejection of claims 1-14. The appellants received an Advisory Action, dated September 1, 2005. The Advisory Action reiterated the rejections of claims 1-14 under the cited prior art.

5. SUMMARY OF INVENTION

The appellant's invention provides a method and apparatus for load balancing by assigning load in a communication system network having multiple servers, each server having a load level based on serving a number of clients.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. (FIG. 5; pages 19-20; pages 25-26)

6

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that.includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2. (FIG. 5; pages 20-21; pages 25-26)

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers. (FIG. 5; page 21; pages 25-26)

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period. (FIG. 5; page 22; pages 25-26)

6. ISSUES

Whether claims 1-4 are unpatentable under 35 U.S.C. §35 U.S.C. §102(e) as being anticipated by Jindal.

7. GROUPING OF CLAIMS

Appellants designate the following group of claims:
Group I: claims 1-14.

8. ARGUMENT

(i) Rejection under 35 U.S.C. §112, first paragraph:

None

(ii) Rejection under 35 U.S.C. §112, second paragraph:

8

None

(iii) Rejection under 35 U.S.C. §102:

The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal. With respect to claim 1, the Examiner stated that Jindal teaches grouping multiple servers into first and second server groups, wherein said first server group has a load level less than a load level of said second server group (col. 9, line 48 to col. 10, line 40) calculating a time period T (col. 2 lines 40-67; col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27), assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; and col. 9, line 14-27), determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group (col. 8, line 55 to col. 10, line 40), and assigning load to a server selected from said another group of servers after expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27).

The appellants respectfully disagree with the Examiner's interpretation of Jindal. Jindal teaches a selection and subsequent re-selection of a server from a same group of servers. That is, Jindal teaches a central server that periodically determines a server from a same group of servers to be the "preferred" server to which client requests are routed. In order to select the preferred server, the central server periodically collects one or more items of information concerning each server in the group, such as a number of clients being serviced by the server at that moment, a number of client requests handled by the server in a predetermined period of time, and a distance of the server from the central server. Based on the collected information, the central server then selects a server from the group of servers to be the preferred server. The group of servers may comprise multiple servers directly connected to the central server (servers 110, 112, and 114 in FIG. 2), or the group of servers may comprise multiple servers that are indirectly connected to the central server (servers 302, 304, 312, and 314, which are connected to central server 100 via intermediate servers 306 and 316 in FIG. 3, that is, the 'server farms' referenced by the Examiner in the Advisory Action). However, in both instances,

9

at each server selection point the preferred server is selected by the central server from the same group of servers. Thus load balancing is performed not by varying the groups of servers analyzed but instead by periodically analyzing each server in a same group of servers.

In the Advisory Action, the Examiner referenced column 10, lines 32-40, of Jindal as teaching the segregation of servers into multiple groups. The appellants respectfully disagree. In this section of Jindal, Jindal merely teaches that the server group may comprise individual, participating servers ('segregated' servers) distributed among separate 'server farms,' rather than all the servers of each of one or more 'farms.' These individual servers may be considered something of a virtual group. Jindal then teaches that each farm, upon receiving a load allocation for the participating servers of that farm and which load allocation is based on a 'global' load distribution policy, may then distribute load to the participating servers of that farm based on a different load distribution policy than the 'global' load distribution policy. However, again, at each selection point, the load is being distributed among the same participating servers, that is, among the members of the virtual group.

By contrast, claim 1 teaches load balancing by varying the groups of servers considered. That is, claim 1 teaches grouping multiple servers into a first and second server groups, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of a time period T, determining another group of servers comprising the group of servers that includes the first server group and further comprising the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. Nowhere does Jindal teach such load balancing. Accordingly, the appellants respectfully submit that claim 1 is not unpatentable over the prior art of record.

Each of claims 2-4 includes limitations of assigning load to a server selected from a group of servers comprising a first server group from an initial time until expiration of said time period T, determining multiple other groups of servers comprising the group of servers that includes the first server group and further comprising another server group of multiple other servers groups, and assigning load to a server selected from a group of

10

servers of the multiple other groups of servers after expiration of said time period T. As noted above, none of these limitations are taught by Jindal. Accordingly, the appellants respectfully submit that claims 2-4 are not unpatentable over the prior art of record.

Regarding dependent claims 5-14, because claims 5-14 depend directly or indirectly from independent claim 4, the appellants respectfully submit that claims 5-14 are not unpatentable over the prior art of record.

(iv) Rejection under 35 U.S.C. §103:

None.

(v) Other rejections

None.

8.   CONCLUSION

For the above reasons, the appellants respectfully submit that the rejection of claims 1-14 under 35 U.S.C. §102(e) as being upatentable over Jindal is in error and should be reversed and the claims allowed.

Respectfully submitted,

David Chen, et al.

By: _____

Steven A. May
Attorney for Appellants
Registration No. 44,912
Tel. No.: 847/576-3635
Fax No.: 847/576-3750

11

APPENDIX

1.   In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a first and second server groups, wherein said first server group has a load level less than a load level of said second server group;

calculating a time period T;

assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T;

determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group; and

assigning load to a server selected from said another group of servers after expiration of said time period T.

2.   In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating time periods T1 and T2, wherein said time period T2 is longer than said time period T1;

assigning load to a server selected from a group of servers comprising said server group G0 from an initial time until expiration of said time period T1;

determining another group of servers comprising said group of servers that includes said server group G0 and further comprising said server group G1;

assigning load to a server selected from said another group of servers after expiration of said time period T1;

determining yet another group of servers comprising said group of servers that includes said another group of servers and further comprising said server group G2; and

12

assigning load to a server selected from said yet another group of servers after expiration of said time period T2.

3. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers.

4. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk corresponding to said server groups G1 through Gk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

13

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers, wherein each group of servers of said other groups of servers corresponds to an expiring time period.

5. The method as recited in claim 4 wherein said plurality of time periods T1 through Tk each is based on a difference between load levels of at least two server groups in said plurality of server groups G0 through Gk.

6. The method as recited in claim 4 further comprising the step of:

receiving an update of load level of at least one of said plurality of servers in said plurality of server groups G0 through Gk;

repeating said grouping to produce a new plurality of server groups G0 through Gk based on said update of load level;

repeating said calculating said plurality of time periods to produce a new plurality of time periods T1 through Tk corresponding to said new plurality of server groups G0 through Gk;

resetting said initial time to a reset initial time, and assigning load to a server selected from servers in said new server group G0 from said reset initial time until expiration of said new time period T1;

assigning load, after expiration of each of said new time periods T1 through Tk measured from said reset initial time, to a server selected from servers in a combination of servers including said new server group G0 and at least one other server group, in said new server groups G1 through Gk, corresponding to an expiring time period.

14

7. The method as recited in claim 4 wherein said grouping of said plurality of server groups G0 through Gk is based on similarity of load levels among said plurality of servers.

8. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a round robin selection method.

9. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a random selection method.

10. The method as recited in claim 4 wherein each of said plurality of time periods T1 through Tk is based on load levels of at least two server groups selected from said plurality of server groups G0 through Gk, a request arrival rate and a server service rate.

11. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of said plurality of servers.

12. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of a combination of servers of said plurality of servers.

13. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of said plurality of servers.

14. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of a combination of servers of said plurality of servers.

15

**RECEIVED
CENTRAL FAX CENTER**

- PATENT -

OCT 1 0 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT: Chen et al.   EXAMINER: Shingles, K.

SERIAL NO.:   09/447,400   ART UNIT: 2141

FILED: 11/22/99   CASE NO.:   CE08051R

ENTITLED: LOAD BALANCING METHOD IN A COMMUNICATION
                 NETWORK.

Motorola, Inc.
Corporate Offices
1303 E. Algonquin Road
Schaumburg, IL 60196
October 10, 2005

APPEAL BRIEF UNDER 37 CFR 1.192

Certificate of Transmission under 37 CFR 1.8

I hereby certify that this correspondence is being
facsimile transmitted to the United States
Patent and Trademark Office.

on _October 10, 2005_

Motorola, Inc.   10/10/2005

Name of applicant, assignee, or   Date
Registered Representative

Signature

Mail Stop Appeal Brief - Patents
Commissioner of Patents
P.O. Box 1450
Alexandria, Va. 22313-1450

Commissioner:

     The appellants hereby respectfully submit the following Appeal Brief in response
to a final Office Action dated May 10, 2005, and a Notice of Appeal filed August 8, 2005.

1.   REAL PARTY IN INTEREST

The real party in interest in this appeal is Motorola, Inc.

2.   RELATED APPEALS AND INTERFERENCES

There are no other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal.

3.   STATUS OF CLAIMS

This is an appeal from a final Office Action, dated May 10, 2005. Claims 1-14 are appealed. In a first Office Action dated August 12, 2002, the Examiner rejected claims 1-4, 7, and 8 under 35 U.S.C. §103(a) as being unpatentable over Dias et al. (U.S. patent no. 6,119,143) in view of Smith (U.S. patent no. 5,835,724). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Adelman et al. (U.S. patent no. 6,006,259). The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Attanasio et al. (U.S. patent no. 5,918,017) and Fine (U.S. patent no. 4,894,846). The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of the applicants' admitted prior art. The Examiner rejected claims 10-14 under 35 U.S.C. §103(a) as being unpatentable over Dias in view of Smith and further in view of Yu (U.S. patent no. 6,078,943). In a response dated November 12, 2002, the appellants replied to the first Office Action without amending the claims.

In a second Office Action dated January 29, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu (U.S. patent no. 6,078,943) in view of Smith. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu view of Smith and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Smith and further in view of

2

the applicants' admitted prior art. In a response dated April 28, 2003, the appellants replied to the second Office Action without amending the claims.

In a third Office Action dated July 14, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt et al. (U.S. patent no. 6,424,992) and further in view of Devarakonda (U.S. patent no. 6,195,680). The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated October 14, 2003, the appellants replied to the third Office Action without amending the claims.

In a fourth Office Action dated December 30, 2003, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and further in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Dias and Devarakonda and further in view of Attanasio and Fine The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Goldszmidt and Devarakonda and further in view of the applicants' admitted prior art. In a response dated June 22, 2004, the appellants replied to the fourth Office Action without amending the claims.

In a fifth Office Action dated October 5, 2004, the Examiner rejected claims 1-4, 7, 8, and 10-14 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda. The Examiner rejected claim 5 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Adelman. The Examiner rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda and further in view of Attanasio and Fine. The Examiner rejected claim 9 under 35 U.S.C. §103(a) as being unpatentable over Yu in view of Devarakonda

3

and further in view of the applicants' admitted prior art. In an Amendment dated February 4, 2005, the appellants amended each of claims 1-4.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T.

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2.

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method

4

including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers.

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period.

In a final Office Action dated May 10, 2005, the Examiner rejected claims 8 and 9 under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claims the subject matter which the applicant regards as the

5

invention, in particular as lacking an antecedent basis for the phrase "said combination." The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal et al. (U.S. patent no. 6,327,622, hereinafter referred to as "Jindal"). The Examiner rejected claims 11-14 under 35 U.S.C. §103(a) as being unpatentable over Jindal in view of Couland et al. (U.S. patent no. 6,253,230). No claims were allowed. The pending claims 1-14 are reproduced below in the attached Appendix.

4. STATUS OF AMENDMENTS

An Amendment and Response to the Final Office Action was filed on August 8, 2005, and is currently pending. In the Response to the Final Office Action, the appellants amended claims 8 and 9 and responded to the Examiner's rejection of claims 1-14. The appellants received an Advisory Action, dated September 1, 2005. The Advisory Action reiterated the rejections of claims 1-14 under the cited prior art.

5. SUMMARY OF INVENTION

The appellant's invention provides a method and apparatus for load balancing by assigning load in a communication system network having multiple servers, each server having a load level based on serving a number of clients.

Claim 1, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into a first and second server groups, wherein the first server group has a load level less than a load level of the second server group, calculating a time period T, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of the time period T, determining another group of servers including the group of servers that includes the first server group and further including the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. (FIG. 5; pages 19-20; pages 25-26)

Claim 2, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple into multiple server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating time periods T1 and T2, wherein the time period T2 is longer than the time period T1, assigning load to a server selected from a group of servers comprising the server group G0 from an initial time until expiration of the time period T1, determining another group of servers comprising the group of servers that includes the server group G0 and further comprising the server group G1, assigning load to a server selected from the another group of servers after expiration of the time period T1, determining yet another group of servers comprising the group of servers that includes the another group of servers and further comprising the server group G2, and assigning load to a server selected from the yet another group of servers after expiration of the time period T2. (FIG. 5; pages 20-21; pages 25-26)

Claim 3, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in the communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers. (FIG. 5; page 21; pages 25-26)

7

Claim 4, as amended, provides a method for assigning load in a communication system network having multiple servers, each of the multiple servers having a load level based on serving a number of clients in said communication system network, the method including grouping the multiple servers into multiple server groups G0 through Gk, wherein the server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level, calculating multiple time periods T1 through Tk corresponding to the server groups G1 through Gk, assigning load to a server selected from servers in the server group G0 from an initial time until expiration of the time period T1, determining multiple other groups of servers, wherein each group of servers of the multiple other groups of servers comprises the group of servers that includes the server group G0 and further comprises at least one other server group selected from the server groups G1 through Gk, and assigning load, after expiration of each of the time periods T1 through Tk measured from the initial time, to a server selected from a group of servers of the multiple other groups of servers, wherein each group of servers of the other groups of servers corresponds to an expiring time period. (FIG. 5; page 22; pages 25-26)

6. ISSUES

Whether claims 1-4 are unpatentable under 35 U.S.C. §35 U.S.C. §102(e) as being anticipated by Jindal.

7. GROUPING OF CLAIMS

Appellants designate the following group of claims:
Group I: claims 1-14.

8. ARGUMENT

(i) Rejection under 35 U.S.C. §112, first paragraph:

None

(ii) Rejection under 35 U.S.C. §112, second paragraph:

8

None

(iii) Rejection under 35 U.S.C. §102:

The Examiner rejected claims 1-10 under 35 U.S.C. §102(e) as being anticipated by Jindal. With respect to claim 1, the Examiner stated that Jindal teaches grouping multiple servers into first and second server groups, wherein said first server group has a load level less than a load level of said second server group (col. 9, line 48 to col. 10, line 40) calculating a time period T (col. 2 lines 40-67; col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27), assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; and col. 9, line 14-27), determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group (col. 8, line 55 to col. 10, line 40), and assigning load to a server selected from said another group of servers after expiration of said time period T (col. 4, lines 57-67; col. 6, lines 35-64; col. 8, lines 1-7 and 47-54; and col. 9, line 14-27).

The appellants respectfully disagree with the Examiner's interpretation of Jindal. Jindal teaches a selection and subsequent re-selection of a server from a same group of servers. That is, Jindal teaches a central server that periodically determines a server from a same group of servers to be the "preferred" server to which client requests are routed. In order to select the preferred server, the central server periodically collects one or more items of information concerning each server in the group, such as a number of clients being serviced by the server at that moment, a number of client requests handled by the server in a predetermined period of time, and a distance of the server from the central server. Based on the collected information, the central server then selects a server from the group of servers to be the preferred server. The group of servers may comprise multiple servers directly connected to the central server (servers 110, 112, and 114 in FIG. 2), or the group of servers may comprise multiple servers that are indirectly connected to the central server (servers 302, 304, 312, and 314, which are connected to central server 100 via intermediate servers 306 and 316 in FIG. 3, that is, the 'server farms' referenced by the Examiner in the Advisory Action). However, in both instances,

9

at each server selection point the preferred server is selected by the central server from the same group of servers. Thus load balancing is performed not by varying the groups of servers analyzed but instead by periodically analyzing each server in a same group of servers.

In the Advisory Action, the Examiner referenced column 10, lines 32-40, of Jindal as teaching the segregation of servers into multiple groups. The appellants respectfully disagree. In this section of Jindal, Jindal merely teaches that the server group may comprise individual, participating servers ('segregated' servers) distributed among separate 'server farms,' rather than all the servers of each of one or more 'farms.' These individual servers may be considered something of a virtual group. Jindal then teaches that each farm, upon receiving a load allocation for the participating servers of that farm and which load allocation is based on a 'global' load distribution policy, may then distribute load to the participating servers of that farm based on a different load distribution policy than the 'global' load distribution policy. However, again, at each selection point, the load is being distributed among the same participating servers, that is, among the members of the virtual group.

By contrast, claim 1 teaches load balancing by varying the groups of servers considered. That is, claim 1 teaches grouping multiple servers into a first and second server groups, assigning load to a server selected from a group of servers comprising the first server group from an initial time until expiration of a time period T, determining another group of servers comprising the group of servers that includes the first server group and further comprising the second server group, and assigning load to a server selected from the another group of servers after expiration of the time period T. Nowhere does Jindal teach such load balancing. Accordingly, the appellants respectfully submit that claim 1 is not unpatentable over the prior art of record.

Each of claims 2-4 includes limitations of assigning load to a server selected from a group of servers comprising a first server group from an initial time until expiration of said time period T, determining multiple other groups of servers comprising the group of servers that includes the first server group and further comprising another server group of multiple other servers groups, and assigning load to a server selected from a group of

10

servers of the multiple other groups of servers after expiration of said time period T. As noted above, none of these limitations are taught by Jindal. Accordingly, the appellants respectfully submit that claims 2-4 are not unpatentable over the prior art of record.

Regarding dependent claims 5-14, because claims 5-14 depend directly or indirectly from independent claim 4, the appellants respectfully submit that claims 5-14 are not unpatentable over the prior art of record.

(iv) Rejection under 35 U.S.C. §103:

None.

(v) Other rejections

None.

8. CONCLUSION

For the above reasons, the appellants respectfully submit that the rejection of claims 1-14 under 35 U.S.C. §102(e) as being upatentable over Jindal is in error and should be reversed and the claims allowed.

Respectfully submitted,

David Chen, et al.

By: _____

Steven A. May
Attorney for Appellants
Registration No. 44,912
Tel. No.: 847/576-3635
Fax No.: 847/576-3750

11

APPENDIX

1. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a first and second server groups, wherein said first server group has a load level less than a load level of said second server group;

calculating a time period T;

assigning load to a server selected from a group of servers comprising said first server group from an initial time until expiration of said time period T;

determining another group of servers comprising said group of servers that includes said first server group and further comprising said second server group; and

assigning load to a server selected from said another group of servers after expiration of said time period T.

2. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through G2, wherein server groups G0 through G2 respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating time periods T1 and T2, wherein said time period T2 is longer than said time period T1;

assigning load to a server selected from a group of servers comprising said server group G0 from an initial time until expiration of said time period T1;

determining another group of servers comprising said group of servers that includes said server group G0 and further comprising said server group G1;

assigning load to a server selected from said another group of servers after expiration of said time period T1;

determining yet another group of servers comprising said group of servers that includes said another group of servers and further comprising said server group G2; and

12

assigning load to a server selected from said yet another group of servers after expiration of said time period T2.

3. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers.

4. In a communication system network having a plurality of servers, each of said plurality of servers having a load level based on serving a number of clients in said communication system network, a method comprising the steps of:

grouping said plurality of servers into a plurality of server groups G0 through Gk, wherein said server groups G0 through Gk respectively have load levels from progressively a least amount of load level to a most amount of load level;

calculating a plurality of time periods T1 through Tk corresponding to said server groups G1 through Gk;

assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1;

13

determining a plurality of other groups of servers, wherein each group of servers of said plurality of other groups of servers comprises said group of servers that includes said server group G0 and further comprises at least one other server group selected from said server groups G1 through Gk; and

assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from a group of servers of the plurality of other groups of servers, wherein each group of servers of said other groups of servers corresponds to an expiring time period.

5. The method as recited in claim 4 wherein said plurality of time periods T1 through Tk each is based on a difference between load levels of at least two server groups in said plurality of server groups G0 through Gk.

6. The method as recited in claim 4 further comprising the step of:

receiving an update of load level of at least one of said plurality of servers in said plurality of server groups G0 through Gk;

repeating said grouping to produce a new plurality of server groups G0 through Gk based on said update of load level;

repeating said calculating said plurality of time periods to produce a new plurality of time periods T1 through Tk corresponding to said new plurality of server groups G0 through Gk;

resetting said initial time to a reset initial time, and assigning load to a server selected from servers in said new server group G0 from said reset initial time until expiration of said new time period T1;

assigning load, after expiration of each of said new time periods T1 through Tk measured from said reset initial time, to a server selected from servers in a combination of servers including said new server group G0 and at least one other server group, in said new server groups G1 through Gk, corresponding to an expiring time period.

14

7. The method as recited in claim 4 wherein said grouping of said plurality of server groups G0 through Gk is based on similarity of load levels among said plurality of servers.

8. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a round robin selection method.

9. The method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from a group of servers of the plurality of other groups of servers is performed according to a random selection method.

10. The method as recited in claim 4 wherein each of said plurality of time periods T1 through Tk is based on load levels of at least two server groups selected from said plurality of server groups G0 through Gk, a request arrival rate and a server service rate.

11. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of said plurality of servers.

12. The method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate of a combination of servers of said plurality of servers.

13. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of said plurality of servers.

14. The method as recited in claim 10 wherein said server service rate is substituted for an average service rate of a combination of servers of said plurality of servers.

15